

# Econometria

## Modelo de Regresión Lineal Simple

Pasquini, Ricardo

IAE Business School  
Universidad Austral

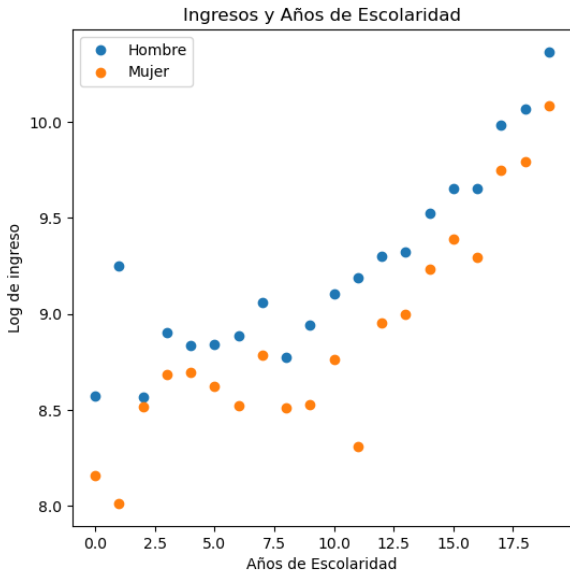
August 19, 2025

# Objetivos de Aprendizaje

- ▶ Definir un modelo de regresión con una variable explicativa.
- ▶ Comprender los componentes del modelo de regresión.
- ▶ Aprender a estimar el modelo de regresión.
- ▶ Interpretar los coeficientes de regresión.
- ▶ Evaluar la bondad de ajuste del modelo.

# Funcion de Esperanza Condicional

Aplicacion: Ingresos y Años de Escolaridad



# ¿Qué es la Regresión de una Variable?

- ▶ Definición de regresión de una variable.
- ▶ Ecuación:  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
- ▶ Explicación de cada término:  $(Y_i), (X_i), (\beta_0), (\beta_1), (\varepsilon_i)$

# Estimación del Modelo de Regresión

- ▶ Método: Mínimos Cuadrados Ordinarios (MCO).
- ▶ Explicación de MCO.
- ▶ Minimización de la suma de residuos cuadrados.

# Estimador de Mínimos Cuadrados Ordinarios (OLS)

- ▶ El estimador OLS es utilizado para estimar los coeficientes de un modelo de regresión lineal simple.
- ▶ Minimiza la suma de los cuadrados de los residuos (RSS), que es la diferencia entre los valores observados de  $Y$  y los valores predichos por el modelo.

# Estimador de Mínimos Cuadrados Ordinarios (OLS)

**Paso 1:** Definir la función de pérdida (RSS):

$$RSS = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

**Paso 2:** Minimizar la función de pérdida en términos de  $\beta_0$  y  $\beta_1$ .

**Paso 3:** Resolver las ecuaciones resultantes para  $\beta_0$  y  $\beta_1$ .

# Estimador de Mínimos Cuadrados Ordinarios (OLS)

## Derivadas Parciales:

Derivamos la función de pérdida (RSS) respecto a  $\beta_0$  y  $\beta_1$  y las igualamos a cero para minimizar RSS:

$$\frac{\partial RSS}{\partial \beta_0} = -2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i) = 0$$

$$\frac{\partial RSS}{\partial \beta_1} = -2 \sum_{i=1}^n X_i (Y_i - \beta_0 - \beta_1 X_i) = 0$$

Estas ecuaciones conducen a las soluciones para  $\hat{\beta}_0$  y  $\hat{\beta}_1$ .



# Estimador de Mínimos Cuadrados Ordinarios (OLS)

Las ecuaciones resultantes para  $\beta_0$  y  $\beta_1$  son:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

Donde  $\bar{X}$  y  $\bar{Y}$  son las medias muestrales de  $X$  y  $Y$  respectivamente.

Estos son los estimadores de mínimos cuadrados ordinarios (OLS) para el modelo de regresión lineal simple.

# Interpretación de los Coeficientes de Regresión

- ▶ Interpretación de  $\beta_0$  y  $\beta_1$ .
- ▶ Cómo interpretar el coeficiente de pendiente ( $\beta_1$ ) en el contexto del modelo.

# Bondad de Ajuste

- ▶ Medidas de bondad de ajuste:  $R^2$ ,  $R^2$  ajustado, estadístico F.
- ▶ Explicación de cada medida y su significado.
- ▶ Interpretación de  $R^2$ .

## Bondad de Ajuste - $R^2$

- ▶ El coeficiente  $R^2$  se define como la proporción de la varianza de la variable dependiente  $Y$  que es explicable por la variable  $X$ .
- ▶ Paso 1: Calcular los valores predichos de  $Y$  utilizando la ecuación de regresión:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Donde  $\hat{Y}_i$  es el valor predicho de  $Y$  para la  $i$ -ésima observación, y  $\hat{\beta}_0$  y  $\hat{\beta}_1$  son los coeficientes estimados obtenidos del análisis de regresión.

- ▶ Paso 2: Calcular la suma total de cuadrados (TSS), que mide la variabilidad total en la variable dependiente:

$$TSS = \sum_{i=1}^n (Y_i - \bar{Y})^2$$

Donde  $\bar{Y}$  es la media de los valores observados de  $Y$ .

## Bondad de Ajuste - $R^2$

- Paso 3: Calcular la suma de cuadrados residual (RSS), que mide la variabilidad que no es explicada por el modelo de regresión:

$$RSS = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

- Paso 4: Calcular la suma de cuadrados explicada (ESS), que mide la variabilidad explicada por el modelo de regresión:

$$ESS = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

- Paso 5: Calcular  $R^2$  utilizando la fórmula:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Esta fórmula representa la proporción de la variabilidad total en  $Y$  que es explicada por el modelo de regresión.

# Error Cuadrático Medio (MSE)

- ▶ El MSE es una medida de la calidad de un modelo de regresión.
- ▶ Indica el promedio de los cuadrados de los errores, es decir, la diferencia entre los valores predichos por el modelo y los valores reales de la variable dependiente.

# Error Cuadrático Medio (MSE)

## Fórmula

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

- ▶ Donde  $Y_i$  es el valor real de la variable dependiente para la  $i$ -ésima observación.
- ▶  $\hat{Y}_i$  es el valor predicho por el modelo de regresión para la  $i$ -ésima observación.
- ▶  $n$  es el número total de observaciones.

# Error Cuadrático Medio (MSE)

- **Paso 1:** Calcular los residuos para cada observación:

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i$$

- **Paso 2:** Elevar al cuadrado cada error:

$$\hat{\varepsilon}_i^2 = (Y_i - \hat{Y}_i)^2$$

- **Paso 3:** Calcular el promedio de los residuos al cuadrado:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n \hat{\varepsilon}_i^2$$



# Ejemplo

- ▶ Conjunto de datos de ejemplo: ingresos y años de educación.
- ▶ Gráfico de dispersión de los datos.
- ▶ Ecuación de regresión:  $Ingresos = \beta_0 + \beta_1 \times Educacion$
- ▶ Interpretación de los coeficientes en el contexto del ejemplo.

# Referencias

- ▶ Wooldridge ch. 2